

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-069471

(43)Date of publication of application : 10.03.1998

(51)Int.Cl.

G06F 15/163
G06F 13/00

(21)Application number : 08-226396

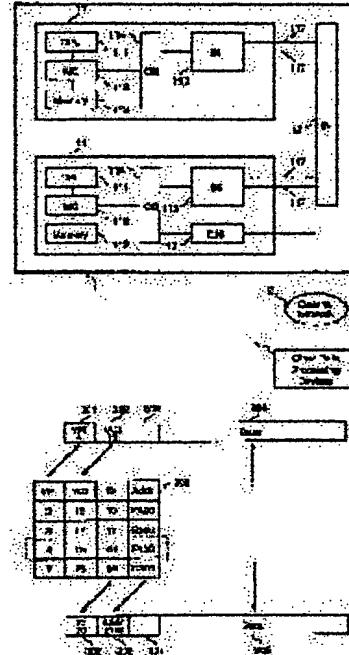
(71)Applicant : HITACHI LTD

(22)Date of filing : 28.08.1996

(72)Inventor : MASHIERU FUREDERIKO
KITAI KATSUYOSHI
YOSHIZAWA SATOSHI
TARUI TOSIYAKI
MURAHASHI HIDEKI**(54) PARALLEL COMPUTER****(57)Abstract:**

PROBLEM TO BE SOLVED: To route a message from an external network to an internal network by providing a gateway to the external network with a correlation table which makes the identifiers of the logical connection of the external network and processes in parallel computers correspond to each other.

SOLUTION: When an asynchronous transfer mode(ATM) cell reaches an external network interface(ENI) 116 from the external network 2, ENI 116 extracts a virtual bus identifier 301 and a virtual channel identifier 302 and performs retrieval from the correlation table 309. Then the corresponding processor identifier IED and buffer address Addr of a destination processor are read out to form messages 305-308 of the internal network 12, and data of an internal network packet 307 are taken directly out of data 304 of an original ATM cell to transmit the generated messages from a gateway processor to the internal network in the parallel computers.

**LEGAL STATUS**

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-69471

(43)公開日 平成10年(1998)3月10日

(51)Int.Cl.⁶
G 0 6 F 15/163
13/00

識別記号 15/163
3 5 5

F I
G 0 6 F 15/16
13/00

技術表示箇所
3 2 0 G
3 5 5

審査請求 未請求 請求項の数5 OL (全8頁)

(21)出願番号

特願平8-226396

(22)出願日

平成8年(1996)8月28日

(71)出願人 000005108

株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72)発明者 マシエル フレデリコ

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 北井 克佳

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 吉澤 聰

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(74)代理人 弁理士 小川 勝男

最終頁に続く

(54)【発明の名称】 並列計算機

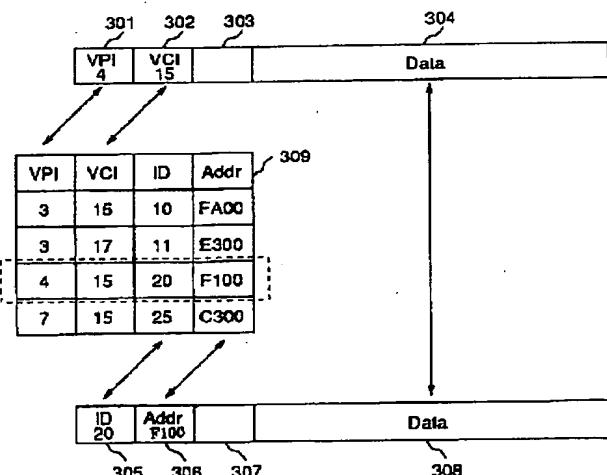
(57)【要約】

【課題】 外部ネットワークENのメッセージMSを並列計算機PPのプロセスへ直接ルーティング可能とし、PPの内部ネットワークINとENを接続するプロセッサの負荷を低減する。

【解決手段】 ENとのゲートウェイGWの働きをするPP内のプロセッサに、ENの論理コネクションの識別子とPP内のプロセス（プロセッサ識別子とプロセッサ内プログラム（プロセス）のアドレス）との対応表と、対応表を参照しプロセスを宛先とするMSをENのMSから作成する手段を設ける。

【効果】 PP内のプロセスとENのMS交換が効率的になる。GWプロセッサのオーバヘッドが減り、PPの処理能力が向上する。

図3



【特許請求の範囲】

【請求項1】複数のプロセッサと内部ネットワークとか
ら構成される並列計算機であって、

前記複数のプロセッサの内の1つのプロセッサは、前記
前記並列計算機に接続される外部ネットワークを接続す
るための外部ネットワークインターフェースを有してお
り、

前記外部ネットワークインターフェースは、前記外部ネッ
トワークでの論理コネクションの識別子と前記並列計算
機内でのプロセスとの対応表と、前記外部ネットワーク
からメッセージを受信したときに前記メッセージ内の論
理コネクション識別子で前記対応表を検索する手段と、
検索手段によって得られた当該論理コネクション識別子
に対応するプロセスを用いて内部ネットワーク用のメ
ッセージを作成する手段とを有し、

前記1つのプロセッサは、前記作成された内部ネットワー
ク用のメッセージを前記内部ネットワークへ送信する
内部ネットワークインターフェースを有することを特徴と
する並列計算機。

【請求項2】前記プロセスは、前記並列計算機内でのプロ
セッサを識別する識別子と当該プロセッサ内のメモリ
のアドレスで規定される請求項1記載の並列計算機。

【請求項3】前記1つのプロセッサは、更に、中央処理
装置と、メモリと、メモリ制御部とを有し、
前記内部ネットワークインターフェースと前記外部ネット
ワークインターフェースが直接通信するための通信手段を
有する請求項1記載の並列計算機。

【請求項4】前記通信手段は、前記メモリ制御部と、前
記外部ネットワークインターフェースと、前記外部ネット
ワークインターフェースを接続する入出力バスである請求
項3記載の並列計算機。

【請求項5】前記メモリ制御部と前記内部ネットワーク
インターフェースが接続され、前記内部ネットワークイン
ターフェースと前記外部ネットワークインターフェースが接
続された請求項4記載の並列計算機。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は並列計算機に関し、
特に、外部ネットワークと接続される並列計算機のイン
ターフェースに関する。

【0002】

【従来の技術】並列コンピュータ、特に大規模並列処理
コンピュータは、内部ネットワークINと呼ぶネットワー
クによって相互接続された多数のプロセッサから構成
されている。図1は、このような並列処理コンピュータ
のアーキテクチャを概略的に示している。並列式マ
シン1には、二つ以上のプロセッサ11が設けられてい
る。各プロセッサ11は、中央処理装置(CPU)11
1と、メモリ制御装置(MC)112と、メモリ113
と、入出力(I/O)バス(LOB)114を有してい

る。図1は、代表的なアーキテクチャを示したものであ
り、図1以外の種々の構成のものがある。プロセッサ1
1は、内部ネットワークインターフェース(INI)11
5によって内部ネットワーク12に接続されている。全
てのプロセッサには、INIが設けられている。

【0003】INI(115)は、内部ネットワーク12との一つ以上
の接続117を有している。接続117の数は、内部ネット
ワーク12のアーキテクチャーに左右される。例えば、n
次元の超立方体網では、各プロセッサ11は、n個の他の
プロセッサ11とn個の接続117を介して接続される。
2次元のメッシュネットワークでは、各プロセッサ11
は、4個の他のプロセッサ11と接続しており、従ってINI
(115)は4個の接続117を有していることになる。

【0004】このようなマシンは、メッセージの形で、
内部ネットワーク12を介してプロセッサ11間のデータ
交換を行うことにより、並行した処理を行っている。
このネットワークに於けるメッセージの転送は、行き先
プロセッサ11を指定してメッセージを送るために、各
プロセッサ11は、全てのプロセッサ11を一意に識別
するための識別子を有している。メッセージ全体は、こ
の行き先プロセッサの識別子とデータから構成されて
いる(尚、メッセージには必要に応じてこれ以外の構成を
有する)。メッセージは、行き先プロセッサの識別子に
応じて内部ネットワーク12内で経路指定される。この
ような経路指定はINI(115)によって行われる。
メッセージは、内部ネットワーク12からINI(11
5)まで送られ、行き先プロセッサに到着するまで転送
(中継転送を含む)される。この送信処理は次のように
行われる。CPU111が内部ネットワーク(INI)
115に命令を送り、メモリ113内のメッセージの位
置を指定する。このメッセージは、MC112とI/O
バス114を介して、メモリ113からINIによって
読み取られ、内部ネットワーク12に送られて、行き先
プロセッサ内のINIによって受け取られる。そして、
MC112とI/Oバス(LOB)114を介して行き
先プロセッサ113のメモリに送られる。メッセージを
送るために用いられるプロトコルは、マシンに応じたも
のである。

【0005】また、並列式マシンは、外部ネットワーク
2を介して他のデータ処理装置5に接続されている。外
部ネットワーク2の例としては、イーサネット、Fib
er Distributed Data Interface(FDDI)や、High-Perfor
mance Parallel Interface(HIPPI)や、Asynchronous Transfe
r Mode(ATM)ネットワーク等が挙げられる。
並列式マシン1のプロセッサ11は、外部ネットワーク
インターフェース(以下、ENIと呼ぶ)116を介して
外部ネットワーク2に接続されている。ここで、全ての

プロセッサが、ENIを有しているわけではない。言い換れば、少なくとも1個以上のプロセッサ（以下、ゲートウェイプロセッサと呼ぶ）がENIを有している。

【0006】外部ネットワーク2と接続されているENI（116）が設けられたゲートウェイプロセッサは、このネットワークを直接アクセスすることができる。他のプロセッサは、内部ネットワーク12とこれらゲートウェイプロセッサを介して外部ネットワーク2にアクセスしなければならない。この処理を以下に詳細に説明する。

【0007】外部ネットワーク2に接続しているENI（116）を持たないプロセッサがこのネットワークにメッセージを送りたい場合、先ず、目標の外部ネットワーク2と接続している外部ネットワークインターフェースENI（116）を有するゲートウェイプロセッサにメッセージを送る。

【0008】ゲートウェイプロセッサは、外部ネットワークへのメッセージを受信したINI（115）による割り込みにより、コンテックストスイッチをオペレーティングシステムに切り替える。オペレーティングシステムは、外部ネットワークへのメッセージを受け取り、それを処理してENI（116）に送る。一方、外部ネットワーク2からのメッセージをENI（116）が受信し、割り込みを発生すると、当該ENIを有するゲートウェイプロセッサは、コンテックストをオペレーティングシステムに切り替える。オペレーティングシステムは、メッセージを受け取り、それを処理して行き先プロセッサを確認する。行き先プロセッサがゲートウェイプロセッサでない場合は、メッセージは、内部ネットワーク12を介して行き先プロセッサのメモリ113内のバッファに送られる。行き先プロセッサが自ゲートウェイプロセッサである場合にはそのまま自ゲートウェイプロセッサ内で処理する。

【0009】外部ネットワーク2のプロトコルにおいては、データは物理ネットワーク上に設定された論理コネクションを介して転送される。このようなネットワークで使用されるプロトコルの例として、インターネットプロトコル（IP）上の伝送制御プロトコル（TCP）およびユーザデータグラムプロトコル（UDP）、インターネットデータグラムプロトコル（IDP）、パケットイクスチェンジプロトコル（PEX）、シーケンスドパケットプロトコル（SPP）、非同期転送モード（ATM）ネットワーク上のプロトコル等が挙げられる。論理コネクションは、各パケットまたはセルのプロトコルヘッダ内のデータによって識別することができる。なお、UDPはコネクションレスプロトコルであるが、プロトコル名と、ソースおよび行き先ネットワークアドレスと、ソースおよび行き先ポートとによって形成される論理結合もここでは論理コネクションと呼ばれて、PEXに適用される。論理コネクションを識別する方法は、使

用されているプロトコルによって異なっている。また、ゲートウェイプロセッサは、外部ネットワーク2からメッセージが送られて来ると、該メッセージが所属する論理コネクションから行き先プロセッサを決定する必要がある。

【0010】このようなゲートウェイでの外部ネットワークと内部ネットワークでの宛先変換の例としては、例えば、特開平3-18961号公報では、並列マシン内で発行される内部メッセージのアドレスが、外部ネットワークであるか並列式マシン内のプロセッサであるかを示すフラグを提案し、並列マシン内でのアドレッシングと、外部ネットワークでのアドレッシングの柔軟性を高めている例が示されている。また、特開平5-181819では、外部ネットワークから、内部ネットワークの行き先プロセッサの行き先プロセッサ番号をデータに添付することで、直接外部ネットワークから並列プロセッサ内のプロセッサへのメッセージへ変換する例が示されている。

【0011】

【発明が解決しようとする課題】並列式マシン1内のプロセッサ11は、効率よくユーザプログラム（プロセス）を実行することに時間を費やすことが望ましい。しかしながら、現状のシステムでは、外部ネットワークからのメッセージは並列計算機内のプロセッサしか指定できず、プロセッサ内のどのプロセス（ユーザプログラム）へのメッセージであるかは、メッセージを受信後、データを解析して初めて知ることができた。このため、各プロセッサではメッセージを受信してから実行までの時間が多くかかっていた。

【0012】また、ゲートウェイプロセッサは、外部ネットワーク2から送られてくる各メッセージによって割り込まれて、オペレーティングシステムがコンテックストスイッチ切り換えをし、メッセージを送信するための経路指定処理を行い、必要な場合には内部ネットワーク12を介して他のプロセッサ11にメッセージを送る処理を行う必要がある。また、外部ネットワーク2に送られるメッセージが内部ネットワーク（12）からゲートウェイプロセッサに送られてきた場合にも、同様の処理が行われる。このような割り込みや処理が行われる結果、ゲートウェイプロセッサのオペレーティングシステムにオーバーヘッドが生じ、ユーザプログラムを実行するのに利用できるゲートウェイプロセッサの時間が減少する。しばしば、ゲートウェイプロセッサはI/Oに独占的に使用され、このため、プロセッサの数を増やすことになりシステムのコストがかさむ結果になる。

【0013】スループット、すなわち所定の時間内に伝送されるデータの量は、ネットワークにとって一つの重要な性能基準である。ゲートウェイプロセッサ内の処理はスループットの向上を制限し、従って性能を低下させる。

【0014】伝送ソース（送信元）から行き先（宛先）までメッセージを送るのに要する時間である待ち時間は、外部ネットワークに関するもう一つの重要な性能基準である。肯定応答メッセージを出す必要があるため、スループットは、コネクション・オリエンティド・プロトコル内の待ち時間に影響される。従って、インターフェースを介したコミュニケーションの全体的性能を上げるために、待ち時間を減らすことはかなり重要なことである。しかし、ゲートウェイ内の処理はコミュニケーションの待ち時間を増加させて、その結果スループットを減少させる。

【0015】なお、特開平3-18961に記載されている方法は、外部ネットワーク内のどのプロセッサへデータを送信するかを示す方法がない。また、特開平5-181819に提案されている方法は、メッセージ内に宛先であるプロセッサ番号を含めなければならず、TCP/IPやATM等の標準ネットワークプロトコルに使用することができない。

【0016】本発明の目的は、外部ネットワークに接続される並列計算機に於いて、外部ネットワークの論理コネクションの識別子に、内部ネットワーク内のプロセッサのプロセスを指定する情報を含めることなく、外部ネットワークからのメッセージを内部ネットワークのプロセッサのプロセスへルーティング可能にすることにある。また、プロセッサのプロセスからのメッセージを外部ネットワークへルーティング可能とすることである。また、上記ルーティングをゲートウェイでの中央処理装置の介入無く上記ルーティングを可能にすることである。

【0017】

【課題を解決するための手段】上記目的を達成するために、複数のプロセッサと内部ネットワークとから構成される並列計算機において、前記複数のプロセッサの内の1つのプロセッサは、前記前記並列計算機に接続される外部ネットワークを接続するための外部ネットワークインターフェースを持ち、前記外部ネットワークインターフェースは、前記外部ネットワークでの論理コネクションの識別子と、前記内部ネットワークでのプロセッサのプロセスとの対応表と、前記外部ネットワークからメッセージを受信したときに前記メッセージ内の論理コネクション識別子で前記対応表を検索する手段と、検索手段によって得られた当該論理コネクション識別子に対応するプロセッサ識別子を用いて内部ネットワーク用のメッセージを作成する手段とを有し、前記1つのプロセッサは、前記作成された内部ネットワーク用のメッセージを前記内部ネットワークへ送信する内部ネットワークインターフェースを有することによって達成される。

【0018】また、前記プロセスは、前記並列計算機内のプロセッサを識別する識別子と当該プロセッサ内のメモリのアドレスで規定することにより達成される。

【0019】また、前記1つのプロセッサは、更に、中央処理装置と、メモリと、メモリ制御部とを有し、前記内部ネットワークインターフェースと前記外部ネットワークインターフェースが直接通信するための通信手段を有することによって達成される。

【0020】次に、前記通信手段は、前記メモリ制御部と、前記外部ネットワークインターフェースと、前記外部ネットワークインターフェースを接続する入出力バスであることによって達成される。

【0021】次に又、前記メモリ制御部と前記内部ネットワークインターフェースが接続され、前記内部ネットワークインターフェースと前記外部ネットワークインターフェースが接続されることによって達成される。

【0022】

【発明の実施の形態】本発明では、プロセッサ11のハードウェア構成に応じて、2通りの形状に形成することができる。第1の構成は、外部ネットワークインターフェースENI(116)をプロセッサ11のI/Oバス114に接続する方法、第2の構成は、プロセッサ11がI/Oバス114を有していない場合に、INI(115)と内部ネットワーク12間の他の接続117と同様の接続118により、外部ネットワークインターフェースENI(116)をINI(115)に接続する方法である。所定の並列式プロセッサ(1)に関するこれら二つの構成の内、どちらが適しているかは、この並列式プロセッサ(1)のハードウェアの特異性によって左右される。これら両方のインターフェース構成を以下に説明する。

【0023】図1は、第1の構成、即ち、外部ネットワークインターフェースENIがゲートウェイプロセッサのI/OバスIOB(114)に付けられている例を示す。論理コネクションの一つに属するパケットまたはセルがENI(116)に到着すると、I/Oバス114を介してINI(115)に直接送られる。このメッセージは、メモリ113からINI(115)に送られるメッセージと同じ方法で、INI(115)によって受け取られる。メッセージは、その後、INI(115)によって内部ネットワーク12に送られて、行き先プロセッサ11に経路指定される。

【0024】図2は、外部ネットワークインターフェースENIが内部ネットワークに付けられている例を示す。論理コネクションの一つに属するパケットまたはセルがENI(116)に到着すると、接続118を介してINI(115)に直接送られる。このメッセージは、内部ネットワーク12から到着するメッセージと同じ方法で、INI(115)によって受け取られる。メッセージは、その後、INI(115)によって行き先プロセッサに経路指定される。

【0025】ENI(116)は、外部ネットワークプロトコルの論理コネクションと、内部ネットワークプロ

トコルのプロセッサ識別子とバッファアドレスとを、対応させる相関テーブルを用いることによって、相互に関連付ける。テーブルのフォーマットと相互関連方法は、外部ネットワーク(2)によって左右される。相互関連方法の例を以下に説明する。

【0026】図3は、ENIが、外部ネットワークである非同期転送モード(ATM)ネットワークのセルと内部ネットワーク用のパケットを変換する例を示す。ATMネットワークの各セル(301、302、303、304)は、ヘッダ(301、302、303)とデータ304から構成され、このヘッダの中には、仮想バス識別子(VPI)301、仮想チャネル識別子(VCI)302、および附加ヘッダデータ303が含まれる。ATMセルが外部ネットワーク2からENI(116)に到着すると、VPI(301)とVCI(302)を用いてセルが属している論理コネクションを識別することができる。

【0027】セルがENIに到着すると、ENIは、VPIおよびVCIを抽出し、これを用いて相関テーブル(309)を検索する。相関テーブル309は、VPI、VCI、行き先プロセッサを識別するためのプロセッサ識別子(ID)、およびバッファアドレス(Address)から構成されている。ここで、バッファアドレスは、プロセッサ内のプロセスを実行するプログラムのプログラムアドレスと1対1に対応するものとする。従って、このアドレスとプロセッサ識別子を組み合わせることで当該並列計算機内のプロセスを一意に特定することができる。

【0028】検索により、対応物(図3において点線のボックスで示されている)が見つかった場合には、行き先プロセッサの対応するプロセッサ識別子(ID)とバッファアドレス(Address)が相関テーブル(309)から読み出されて、ヘッダ(305、306、307)とデータ(308)から構成される内部ネットワーク(12)メッセージ(305、306、307、308)が形成される。

【0029】ヘッダ(305、306、307)は、行き先プロセッサの識別子(305)、行き先プロセッサのバッファアドレス(306)、そして内部ネットワーク(12)に依存している追加データ(307)を含む。また、オリジナルATMセル(301、302、303、304)の追加ヘッダデータ(303)を含んでもよい。内部ネットワークパケット(307)のデータは、オリジナルATMセルのデータ304から直接取り出される。

【0030】作成されたメッセージは、ゲートウェイプロセッサから並列計算機内の内部ネットワークへ送信される。行き先プロセッサが有するプロセッサIDとこのメッセージ内のプロセッサIDが一致したことで、行き先プロセッサがこのメッセージを受信すると、メッセー

ジ内のバッファアドレスを参照してバッファアドレスに当該メッセージが格納される。このバッファアドレスは、プロセスのプログラムアドレスと対応しているので、当該プログラムは直ちにバッファ内のメッセージを処理することができる。

【0031】並列計算機内のプロセスが外部ネットワークへメッセージを送信する際は、まず、ゲートウェイへ内部ネットワーク12からメッセージ(305、306、307、308)がENI(116)に到着した時には、この内部ネットワークからのメッセージをATMセル(301、302、303、304)に変換させるために、逆の処理、即ち、IDおよびAddressを元に相関テーブル309が検索され、VPIおよびVCIが決定され、ATMセルが作成される。なお、この時用いられるIDおよびAddressは、この内部ネットワークでのメッセージの送信元(プロセス)を識別するためのIDおよびAddressである。

【0032】図4は、ENIが、外部ネットワークであるイーサネット、FDDI、HIPPI等のパケット交換ネットワーク、IP、IDP等のネットワーク層プロトコル、UDP、TCP、PEX、SPP等の伝送層プロトコルのパケットと内部ネットワーク用のパケットを変換する例を示す。以下の説明において、TCP/IPおよびUDP/IP用語を用いているが、この両方に記述されているプロトコルの性質は、XNSプロトコルにも適用することができる。

【0033】各外部ネットワーク2のパケット(401、402、403、404)は、物理層ヘッダ(PLH)401、ネットワーク層ヘッダ(NLH)402、伝送層ヘッダ(TLH)403およびデータ404を含んでいる。PLH(401)は、当該パケットのネットワーク層プロトコル(NLP)が何であるかを示す情報を含んでいる。NLH(402)は、ローカルアドレスNAL、遠隔ネットワークアドレスNAR、および、パケットの伝送層プロトコル(TLP)が何であるかを示す情報を含んでいる。TLH(403)は、論理コネクションのローカルポート番号PLおよび遠隔ポート番号PRを含んでいる。ENI(116)は、これらNLP、TLP、NAL、NAR、PL、PRからENI(116)を用いて論理コネクションを識別する。ENIは、外部ネットワーク(2)からパケットが到着した時、パケットからNLP、TLP、NAL、NAR、PL、PRを抽出し、相関テーブル(409)を検索する。

【0034】相関テーブル409は、上記NLP、TLP、NAL、NAR、PL、PRと、内部ネットワークでのプロセッサ識別子IDとバッファアドレスAddressで構成される。ここで、バッファアドレスは、プロセッサ内のプロセスを実行するプログラムのプログラムアド

レスと1対1に対応するものとする。従って、このバッファアドレスとプロセッサ識別子を組み合わせることで当該並列計算機内のプロセスを一意に特定することができる。

【0035】検索により、対応物（図4において、点線のボックスで示されている）が見つかった場合には、行き先プロセッサの対応するプロセッサ識別子（ID）とバッファアドレス（Add r）が相關テーブル（409）から読み出されて、内部ネットワーク（12）でのメッセージ（305、306、307、308）が形成される。追加データ307には、PLH（401）、NLH（402）、またはTLH（403）のデータが必要に応じて含まれる。この形成されたメッセージは、内部ネットワークへ送信されて、行き先プロセッサにて受信され処理される。

【0036】一方、内部ネットワーク12のメッセージ（305、306、307、308）が内部ネットワーク（12）からENI（116）に到着した時に、この内部ネットワーク（12）からのメッセージ（305、306、307、308）を外部ネットワーク（2）パケット（401、402、403、404）に変換するために、逆の処理が行われる。

【0037】上記の両方のタイプのネットワークでは、セル（301、302、303、304）またはパケット（401、402、403、404）が外部ネットワーク（2）から到着した時に、相關テーブル（309または409）に対応物が見つからない場合、これらセル（301、302、303、304）またはパケット（401、402、403、404）は現在の技術で行われているようにゲートウェイプロセッサのオペレーティングシステムによって受け取られる。また、内部ネットワークメッセージ（305、306、307、308）がENI（116）に到着して、相關テーブル（309または409）に対応物が見つからない場合には、ENI（116）は並列式マシン（1）のオペレーティングシステムにエラーを示す。

【0038】ここで、相關テーブルの設定について説明する。相關テーブル（309）エントリの追加と削除は、並列式マシン1のオペレーティングシステムによって、行われる。

【0039】コネクション向きプロトコルの場合、コネクション確立の時にエントリを追加する。コネクションレスプロトコルの場合、2つのケースにエントリを追加する。1つ目は、並列式マシン1で実行しているプログラムが繰り返して送信に使用するアドレスを指定するケースである（例えば、では、Connect関数を呼び出す時）。2つ目は、並列式マシン1で実行しているプログラムがメッセージを受信に使用するためのアドレス（NALとPL）を指定するケースである（例えば、Berkeley Sockets APIでは、bin

d関数を呼び出す時）。この2つ目のケースには、追加するエントリにはローカルアドレスとしてワイルドカードアドレスを使用する。そのワイルドカードアドレスは、実際のネットワークアドレスに対応せず（例えば、IPアドレス0.0.0.0、TCPポート0とUDPポート0）、どんなリモートアドレスでもそのエントリに対応することを意味する。この場合、本発明は受信のみに使用し、送信を現在の技術通りに行う。全てのプロトコルでは、コネクションを解除する時に相關テーブル（309、409）からエントリを削除する。

【0040】図5は、上記のパケットやセルの変換処理相関処理を可能にするENI（116）のハードウェア構成を示す。ENIは、ENI全体を制御するインターフェースCPU1161と、EMIを制御するためのプログラムやデータを記録するためのインターフェースメモリ（IF Mem）1162と、INIとのインターフェースである内部ネットワーク接続インターフェース（INC）1163またはIOB114とのインターフェースであるバス接続インターフェース（BC）1164と、外部ネットワーク2と接続するための外部接続インターフェース（EC）1165と、これらを全て接続するバス（1166）によって構成される。図5にはINC（1163）およびBC（1164）の両方が示されているが、ENIは、図1の形態で用いられる場合はBCのみを有し、図2の形態で用いる場合にはINCのみを有する形で用いて良い。

【0041】セル（301、302、303、304）またはパケット（401、402、403、404）が外部ネットワーク2から送られると、EC1165によって受け取られインターフェースメモリ1162に入れられる。インターフェースCPU1161はECより割り込まれてこのセル（301、302、303、304）またはパケット（401、402、403、404）の到着を示す。その後、インターフェースCPUは、図3または4に示されるようなプロトコル相関処理を行い、内部ネットワークメッセージ（305、306、307、308）を得る。

【0042】ENI（116）が図1に示したようにI/Oバス114に接続されている例では、結果として得られた内部ネットワークメッセージ（305、306、307、308）は、プロセッサメモリ113のようにCPU111メモリスペース内でマップされたインターフェースメモリ1162内に保持される。インターフェースCPU1161は、その後、BC1164およびI/Oバス114を介して、外部ネットワークから受信したメッセージの処理を依頼する命令をメモリ1162内でのメッセージが格納されているアドレスと共にINI115に送り、インターフェースメモリ1162内での内部ネットワークメッセージ（305、306、307、308）の位置を特定する。次に、INI（115）はこの

パケット（メッセージ）を直接ENI内のメモリから、メッセージアドレスを元に読み取り、行き先プロセッサIDを元に、それを内部ネットワーク（12）を介して行き先プロセッサに送る。

【0043】ENI（116）が図2に示したようにINI（115）に接続されている例では、内部ネットワークメッセージ（305、306、307、308）は直接、INI（115）に送られて、行き先プロセッサに経路指定される。内部ネットワークメッセージ（305、306、307、308）が内部ネットワーク（12）から受け取られた場合には、それを外部ネットワーク（2）に送るために逆の処理が行われる。

【0044】

【発明の効果】本発明は、外部ネットワークの論理コネクションの識別子を並列計算機内のプロセス（プロセッサ識別子とプロセッサ内のプログラム（プロセス）のアドレス）とを対応づける相関テーブルを、並列計算機内の外部ネットワークとのゲートウェイの働きをするプロセサに設けたので、並列計算機でのプロセスの実行部を宛先とするメッセージを外部ネットワークのメッセージから直接作成することができ、並列計算機内のプロセスと外部ネットワークのメッセージ交換が効率的に実現できる。

【0045】また、上記ゲートウェイの働きをするプロセッサでは、当該ゲートウェイプロセッサの中央処理装置を用いずに、並列計算機でのプロセスの実行部を宛先とするメッセージを外部ネットワークのメッセージから直接作成するので、ゲートウェイプロセッサでの中央処

理装置のオーバヘッドがへり、並列計算機全体としての処理能力が向上する効果が得られる。

【図面の簡単な説明】

【図1】外部ネットワークインターフェースがI/Oバスに付けられている構成を示した図である。

【図2】外部ネットワークインターフェースが内部ネットワークインターフェースに付けられている構成を示した図である。

【図3】ATMセルと内部ネットワークメッセージの変換を示した図である。

【図4】イーサネット、FDDI、HIPPIのパケットと内部ネットワークメッセージの変換を示した図である。

【図5】ENIのハードウェア構成を示した図である。

【符号の説明】

1 並列式マシン、

2 外部ネットワーク、

11 プロセッサ、

12 内部ネットワーク（INI）、

111 中央処理装置（CPU）、

112 メモリ、

113 入出力バス（IOB）、

114 INI、

115 ENI、

116 ATMセル、

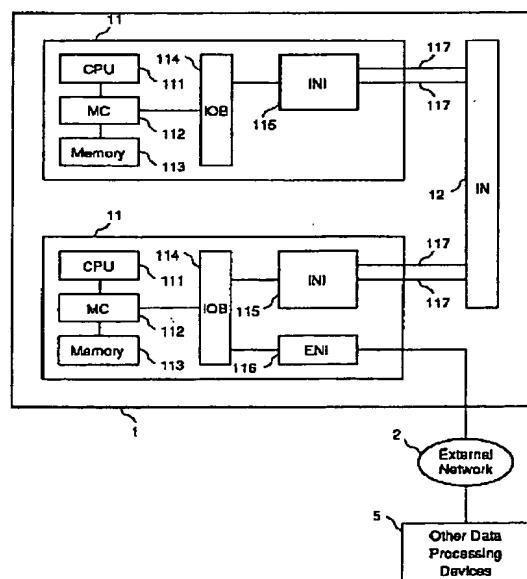
301～304 内部ネットワークメッセージ、

305～308 相関テーブル、

401～404 外部ネットワークパケット。

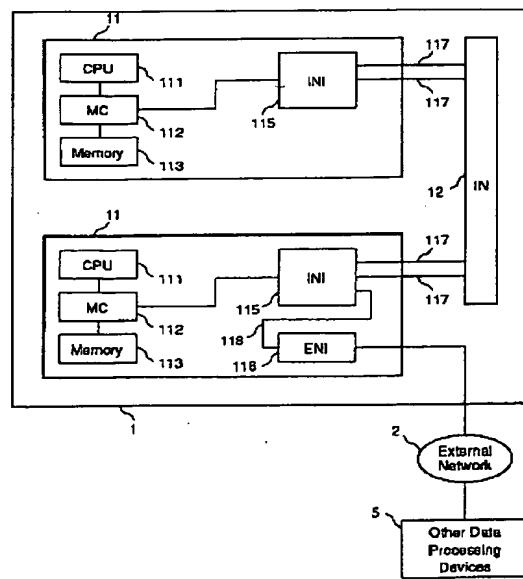
【図1】

図1



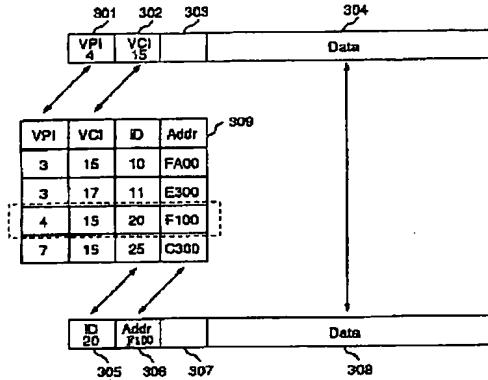
【図2】

図2



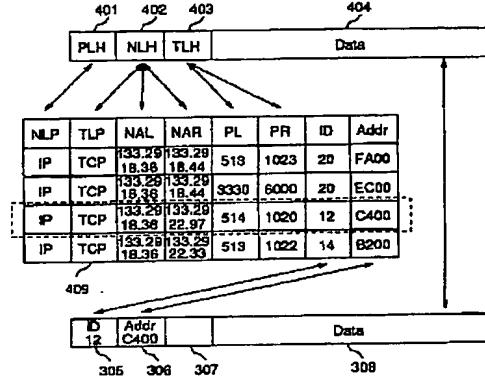
【図3】

図3



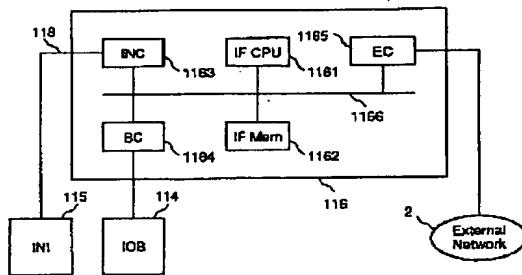
【図4】

図4



【図5】

図5



フロントページの続き

(72)発明者 垂井 俊明

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 村橋 英樹

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内